University of Alberta

DNA POLYMERASE CATALYZED EXPANSION MUTAGENESIS *IN VITRO*

By

Adrian Dobrowsky

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Molecular Biology and Genetics

Department of Biological Sciences

Edmonton, Alberta

Spring 1999

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-40041-7

Canada

University of Alberta

Library Release Form

Name of Author: Adrian Dobrowsky

Title of Thesis: DNA polymerase catalyzed expansion mutagenesis *in vitro*

Degree: Master of Science

Year this Degree Granted: 1999

Permission is hereby granted to the University of Alberta to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

A Dobrowsky

7 Cambria Court
Hamilton, ON
L9C 3Z2

Date: March 1/99

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled DNA POLYMERASE CATALYZED EXPANSION MUTAGENESIS *IN VITRO* submitted by Adrian Dobrowsky in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

Linda J. Reha-Krantz (Supervisor)

Michael Weinfeld

John Bell

Date: March 3, 1994

ABSTRACT

DNA replication usually occurs with high fidelity, however, certain 'slippery' regions of DNA are hotspots for mutation. These errors arise in DNA sequences containing runs of mono-, di-, and trinucleotide repeats. Repeat instability is characteristic of many cancer cells and several neurodegenerative diseases. To improve our understanding of the role of DNA polymerase in repeat expansion, an *in vitro* dinucleotide expansion assay was developed and characterized. The assay has significant improvements over previous assays, including a DNA substrate design that mimics repeat sequences in a chromosomal location. Wild type DNA polymerase was found to catalyze dinucleotide expansion with this DNA substrate. A further study of functionally distinct T4 DNA polymerase mutants suggests that DNA polymerases catalyze strand misalignment and repeat expansion during exonucleolytic proofreading.

ACKNOWLEDGEMENTS

I sincerely thank my supervisor, Dr. Linda Reha-Krantz for her guidance and assistance throughout my master's degree. I also appreciate the helpful advice and suggestions of my committee members, Dr. Mike Weinfeld and Dr. John Bell. Laboratory colleagues, Dr. Koji Hiratsuka and Dr. Mandel provided valuable instruction and technical assistance.

TABLE OF CONTENTS

INTRODUCTION.....1

EXPERIMENTAL PROCEDURES.....16

DNA Polymerases.....16

Synthesis and Purification of DNA Oligonucleotides.....16

Removal of the S_p isomer from Phosphorothioate Oligonucleotides.....18

End Labeling and Annealing Reactions.....18

Dinucleotide Expansion Assay.....18

Maxam-Gilbert Sequencing of Expansion Products.....19

RESULTS.....20

In vitro dinucleotide expansion assay.....20

Sequence of expansion products.....20

Role of the downstream oligonucleotide and DNA polymerase proofreading.....23

Phosphorothioates to block 3'→5' exonuclease activity.....28

Importance of Strand Separation.....34

DISCUSSION.....38

Characterization of the in vitro dinucleotide expansion assay.....38

Probing the importance of proofreading with T4 DNA polymerase mutants.....41

Summary and future directions.....44

REFERENCES.....47

LIST OF TABLES

Table	Description	Page
I.	Characterization of the T4 DNA Polymerase Mutants Used in the Dinucleotide Expansion Assay	12
II.	DNA Oligonucleotides Synthesized for the Dinucleotide Expansion Assay	17

LIST OF FIGURES

Figure	Description	Page
1.	Model for misalignment mutagenesis	2
2.	Mechanism for the inactivation of the human APC gene	3
3.	M13mp2 mutagenesis assay	5
4.	Expansion of repeat sequences as proposed by Schlötterer and Tautz (1992)	8
5.	A minimal scheme for T4 DNA polymerase proofreading	10
6..	DNA template-primer substrate	15
7.	Dinucleotide expansion catalyzed by wild type T4 DNA polymerase	22
8.	Determination of expansion product composition	25
9.	Determination of expansion product sequence	27
10.	Determination of the importance of the downstream oligonucleotide and DNA polymerase proofreading for dinucleotide expansion	30
11.	Placement of phosphorothioates on the DNA substrate	31
12.	Using phosphorothioates to determine the importance 3'→5' exonucleolytic proofreading	33
13.	Comparison of T4 DNA polymerase mutants' ability to promote dinucleotide expansion	36
14.	Model for strand separation and misalignment of CA/GT repeats during proofreading	40
15.	The circular CA/GT repeat containing DNA substrate	46

LIST OF ABBREVIATIONS

$(NH_4)_2SO_4$	ammonium sulfate
A	alanine
APC	adenoma polyposis cancer
D	aspartate
dGTP	deoxyguanosine triphosphate
DMS	dimethyl sulfate
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
dRpase	deoxyribosephosphodiesterase
dTTP	deoxythymidine triphosphate
E	glutamate
EDTA	ethylenediaminetetraacetic acid
G	glycine
HNPCC	hereditary non-polyposis colon cancer
HZ	hydrazine
$KMnO_4$	potassium permanganate
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulfate
MLH1	MutL homolog 1
MMR	mismatch repair
MSH1	MutS homolog 1
MSH3	MutS homolog 3
NaCl	sodium chloride
PCNA	proliferating cell nuclear antigen
PMS1	post meiotic segregation 1
RT	Reverse Transcriptase
S	serine
Taq	*Thermus aquaticus*
Tris-HCl	tris hydroxymethyl aminomethane hydrochloride
tRNA	transfer ribonucleic acid
UV	ultraviolet
W	tryptophan

LIST OF SYMBOLS

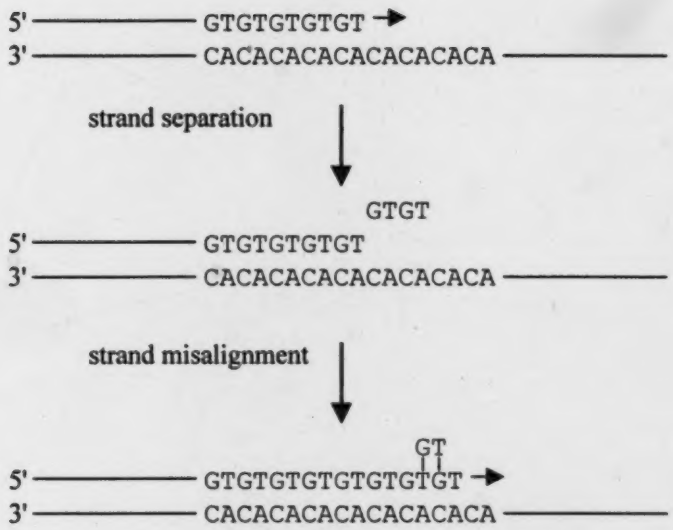
α	alpha
β	beta
δ	delta

INTRODUCTION

Replication of DNA sequences within tracts of mono-, di-, and trinucleotide repeats is error prone (Streisinger *et al.*, 1966; Kunkel, 1985a&b; Strand *et al.*, 1993). Streisinger *et al.* (1966) proposed that repeat length instability is due to the separation and realignment of the primer and template strands producing an extrahelical loop in one of the DNA strands during DNA replication (Fig 1). If replication resumes before the correct strand alignment is restored, the number of repeats is either expanded or contracted, depending on whether the loop is in the primer or template strand, respectively. This can result in frameshift mutations or genome wide instability, *in vivo*. Repeat length instability also depends on the number of repeats. As the number of repeats increases there is a greater probability of forming a misaligned DNA substrate (Streisinger *et al.*, 1966; Kunkel, 1990).

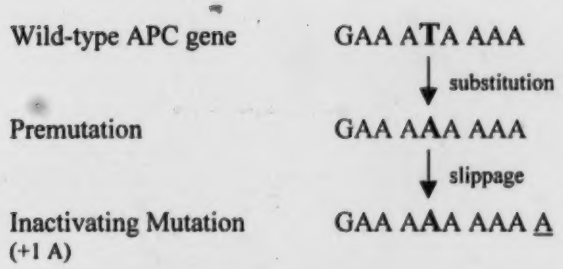
Human disease is often associated with repeat instability. Recently, a mutational mechanism involving the human adenomatous polyposis coli (APC) gene was identified, which links DNA replication and misalignment mutagenesis. The human APC gene normally holds cellular growth in check, but when mutated, the cells become predisposed to colon cancer (Laken *et al.*, 1997). The wild type gene contains a sequence of seven adenine (A) residues that is interrupted by a single thymine (T) residue (see Fig 2). The sequence in adenomatous polyposis cancer patients is a monotonic run of eight adenine (A) residues. The initial mutation is a single base substitution of adenine for the thymine residue. This mutation does not directly affect the function of the gene; rather, it creates a 'hot-spot' for misalignment errors during DNA replication, which inactivate gene function and increase the individual's risk of acquiring colon cancer.

Repeat instability is also a marker for another type of colon cancer, hereditary non-polyposis colon cancer (HNPCC). Microsatellite instability is detected in tumors from individuals diagnosed with HNPCC (Peinada *et al.*, 1992; Ionov *et al.*, 1993). Strand *et al.* (1993) demonstrated that *Saccharomyces cerevisiae* strains defective in mismatch repair (MMR) exhibit microsatellite instability. Tumor cells from HNPCC patients lack MMR activity due to inactivating mutations in the human MMR genes, MSH1, MSH3, PMS1, and MLH1 (Strand *et al.*, 1993). HNPCC patients inherit one



(Adapted from Streisinger *et al.*, 1966)

Figure 1. **Model for misalignment mutagenesis**. During DNA replication in the region containing repeat sequences, the DNA polymerase 'pauses' and dissociates. The DNA strands may separate and reassociate in a misaligned conformation. Misalignment in the primer strand gives rise to expansion of the repeat tract while misalignment of the template strand results in a contraction of the number of repeats. The example above illustrates misalignment of the primer strand.



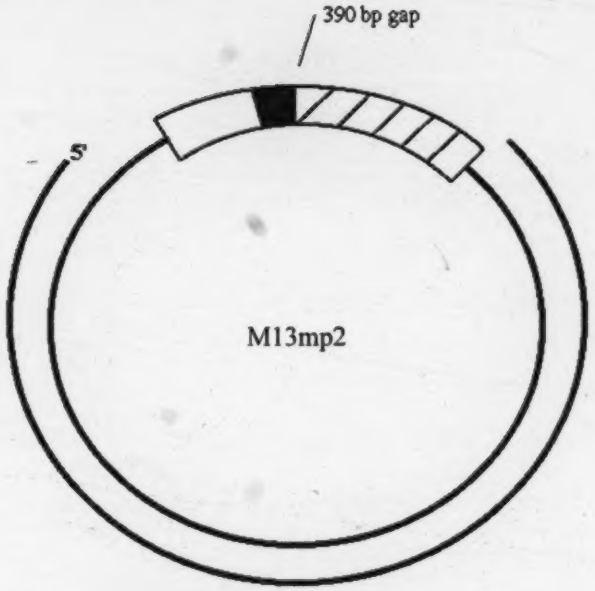
(Adapted from Gura, 1997)

Figure 2. **Mechanism for the inactivation of the human APC gene**. The wild-type APC gene contains a sequence of seven adenines interrupted by a single thymine. This gene is normally responsible for keeping cellular growth in check. The pre-mutation is a substitution of adenine for thymine. APC patients inherit a mutant gene copy from each parent. This mutation does not affect the gene product, however, the run of adenines is a 'hot-spot' for mutations produced by DNA strand misalignments (Laken *et al.*, 1997). Mutations that inactivate APC gene function promote uncontrolled cell growth.

inactive MMR allele and a second inactivating mutation arises in the second copy in pre-tumor cells. The loss of both gene copies leads to increased microsatellite instability throughout the genome, and tumorigenesis occurs. Mismatch repair acts in DNA replication to correct DNA polymerase errors (reviewed by Modrich, 1996) and to prevent homeologous recombination (Whitehouse *et al.*, 1963). Thus, microsatellite instability may occur during either DNA replication or recombination. Reduction of homologous recombination in *Saccharomyces cerevisiae* however, does not decrease the amount of microsatellite instability (Strand *et al.*, 1993). In the absence of post-replicative MMR, strand misalignment during DNA replication is not corrected, resulting in the gain or loss of repeat sequences.

Instability in sequences of trinucleotide repeats can also produce human disease. Fragile X syndrome is one such example. In Fragile X (A) patients, the CGG/GCC repeat at the fragile X mental retardation (FRAXA) locus is expanded from 6 to 54 copies to more than 200. The variability in the number of repeats in unaffected individuals indicates inherent instability. Once the number of repeats reaches a certain length (in this case 52-200 repeats), known as the 'anticipatory' stage, a large expansion is detected and disease results. This pattern of variability in triplet length leading to anticipation and a large expansion is detected for a number of other diseases (see review by Ashley and Warren, 1995).

The instability of repeat sequences has been investigated in *in vitro* assays (Kunkel, 1985a&b). Mammalian DNA polymerases beta (β), alpha (α), and delta (δ) were used to replicate the DNA substrate shown in Figure 3. Bacteriophage M13mp2 DNA containing a 390 base pair gap was used in a forward mutation assay to determine the frequency and specificity of mutation produced *in vitro* by DNA polymerases β, α, and δ. DNA polymerase β is a repair enzyme while α and δ are replicative DNA polymerases. Three classes of errors were identified, including frameshifts, single base substitutions, and deletions. Frameshift mutations were produced in 'hot-spots' consisting of runs of identical bases implicating misalignment mutagenesis during DNA replication. Frequencies of frameshifts were highest for DNA polymerase β and lowest for DNA polymerase δ.



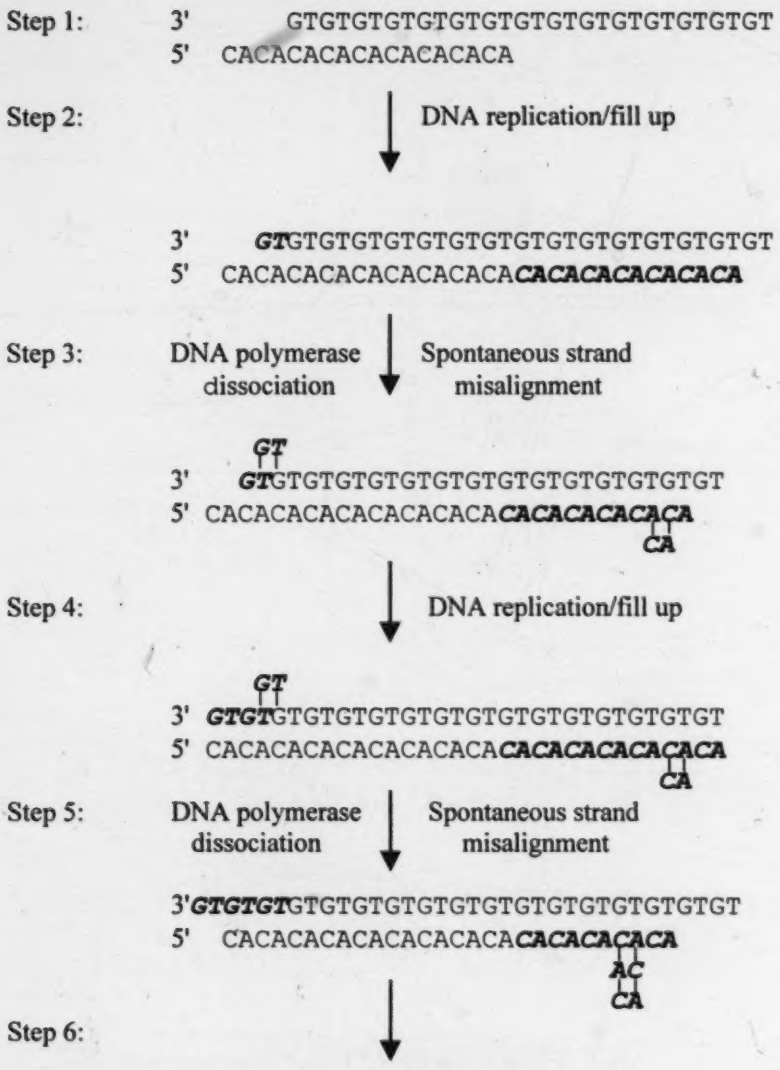
(Adapted from Kunkel, 1985)

Figure 3. **M13mp2 mutagenesis assay**. The *in vitro* forward mutation assay used bacteriophage M13mp2 DNA that contained the C-terminal sequence of the *Escherichia coli* lacI gene (open region), lac promoter and operator regions (solid region), and the N-terminal sequence of the lacZ gene (striped region). A specific gap of 390 base pairs (bp) was created by cleavage with restriction endonucleases. Replication of the gap was conducted with DNA polymerase β, α, and δ. The fidelity of *in vitro* DNA synthesis is determined using the target lacZα gene, by scoring for any error causing loss of a non-essential gene function (α-complementation). Two mutational 'hot-spots' for frameshift mutations were identified. The frameshifts occurred in runs of identical bases implicating misalignment mutagenesis during DNA replication.

Schlötterer and Tautz (1992) also tested the Streisinger slippage model (1966) with an *in vitro* assay. Expansion of several types of di- and trinucleotide repeats was demonstrated. Expansion of GT and CA repeats is illustrated in Figure 4. A short $(CA)_9$ oligonucleotide was annealed to a $(GT)_{15}$ oligonucleotide. Annealing is non-specific so a variety of duplex DNA substrates are produced. Repeat expansion is proposed to be due to a series of DNA replication, DNA polymerase dissociation, and DNA misalignment steps (Fig 4). When the DNA polymerase approaches the end of the template, an unstable DNA-enzyme complex is formed, resulting in the dissociation of the DNA polymerase. The DNA strands realign spontaneously, to create new overhanging 5'-ends. DNA polymerase is then able to reassociate and continue replication. As the end of the template is reached again, an unstable DNA-enzyme complex forms, and the cycle continues. The DNA substrates expanded continuously until the reaction components were exhausted.

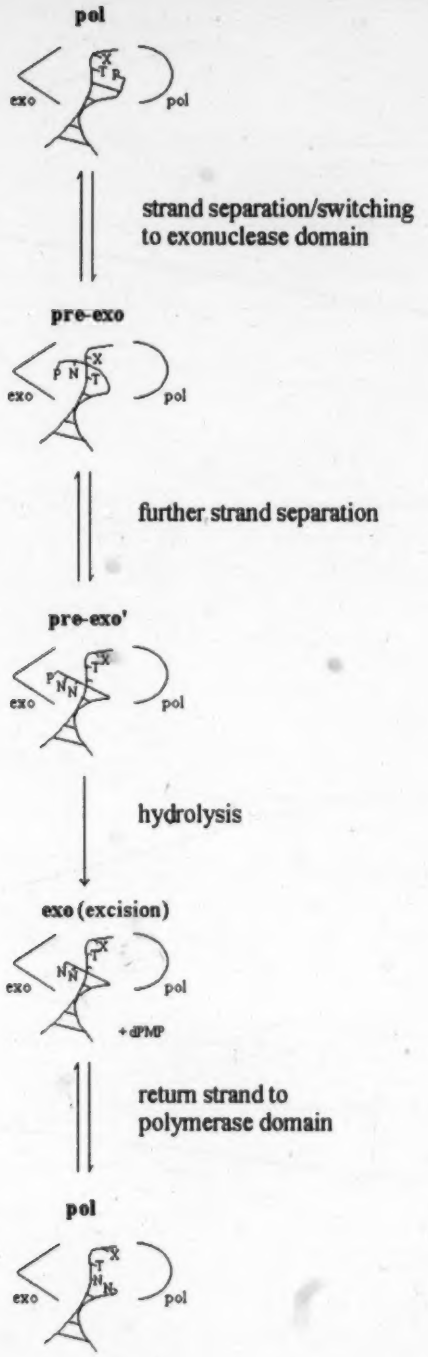
Although the misalignment of repetitive DNA sequences occurs during DNA replication, the above studies do not address whether DNA strand misalignment occurs spontaneously or if the DNA polymerase plays an active role in the mechanism. DNA polymerases that proofread DNA have an opportunity to promote strand misalignment. Exonucleolytic proofreading involves a number of steps and a minimal scheme is outlined in Figure 5. Partial melting of the primer-terminus allows for the transfer of the primer strand from the polymerase active centre to the pre-exonuclease complex. Further strand separation permits binding of the separated primer strand into the exonuclease active centre. The hydrolysis reaction follows, releasing the terminal deoxynucleoside monophosphate. Finally, the primer strand is returned to the polymerase active centre (Baker and Reha-Krantz, 1998). Thus, a mismatched primer-terminus alone is not sufficient to promote the proofreading reaction, but requires DNA polymerase activity in strand separation and DNA transfer between the polymerase and exonuclease active centres. This active participation of the DNA polymerase could occasionally produce strand misalignment. The opportunity for strand misalignment is enhanced in regions of repetitive DNA sequences. This is due to the difficulty that DNA polymerases have replicating repetitive DNA. DNA polymerases "pause" in these areas during replication and this

Figure 4. **Expansion of repeat sequences as proposed by Schlötterer and Tautz (1992).** Step 1: Initial non-specific annealing of repeat oligonucleotides which generates overhanging 5'-ends. Step 2: DNA replication. DNA polymerase incorporates nucleotides (bold-italics) until the ends of the substrate are approached. The DNA polymerase is unstable at the ends of the substrate and dissociates from the DNA substrate. Step 3: The DNA strands slip and create a transient loop. Step 4: Replication of the misaligned DNA strands. Step 5: DNA polymerase dissociation and further strand misalignment. Step 6: Continued cycles of strand misalignment and replication.



(Adapted from Schlötterer and Tautz, 1992)

Figure 5. **A minimal scheme for T4 DNA polymerase proofreading.** P corresponds to the nucleotide at the primer-terminus. The " < " and ") " shaped symbols represent the exonuclease and polymerase active centres, respectively. Exonucleolytic proofreading requires a separation of the DNA strands, followed by switching of the primer strand to the exonuclease active centre, further strand separation, excision of the terminal nucleotide, and the return of the primer strand to the polymerase active centre.



(Adapted from Baker and Reha-Krantz, 1998)

is indicative of DNA polymerase "idling". "Idling" occurs when DNA polymerases go through futile cycles of proofreading and synthesis at a particular site in the DNA sequence. Thus, increased "idling" is a key component in providing the opportunity for strand misalignment. The subject of my thesis is to determine if DNA polymerases play an active role in misalignment mutagenesis. Although I will be testing this proposal with T4 DNA polymerase, which catalyzes a potent exonucleolytic proofreading reaction, the proposal that DNA polymerases play an active role in strand misalignment can be extended to all polymerases that have more than one DNA binding site. These include human DNA polymerase beta (β), which contains a 5'→3' DNA polymerase activity and a 5'-deoxyribosephosphodiesterase (dRpase) activity (Matsumoto and Kim, 1995), and reverse transcriptase (RT) which contains both an RNA-directed DNA polymerase activity and an RNaseH activity (Mölling *et al.*, 1971).

Bacteriophage T4 DNA polymerase was the model DNA polymerase in my studies because of its extensive genetic and biochemical characterization. The DNA replication complex, which consists of DNA polymerase, clamp protein, single-stranded DNA binding protein, primase, helicase, and complex loading proteins has been identified and the genes encoding the constituents have all been cloned for *in vitro* assembly (see review by Nossal, 1994 and references therein). An extensive genetic characterization of T4 DNA polymerase has also provided a variety of mutants to probe DNA polymerase function. Mutator and antimutator DNA polymerases have been identified (Speyer *et al.*, 1966; Drake *et al.*, 1969), which have provided key insights into the biochemical basis of exonucleolytic proofreading (Muzyczka *et al.*, 1972). Mutant DNA polymerases that affect specific steps of the proofreading pathway, from initiation (Reha-Krantz and Nonay, 1994) to strand separation and formation of a pre-exonuclease complex (Stocki *et al.*, 1995; Marquez and Reha-Krantz, 1996; Baker and Reha-Krantz, 1998; Reha-Krantz *et al.*, 1998), and the hydrolysis reaction (Reha-Krantz and Nonay, 1993; Frey *et al.*, 1993), have been identified. Mutant DNA polymerases defective in strand separation and transfer of the strands from polymerase to exonuclease active centre are important for my studies as explained below.

Table I. **Characterization of the T4 DNA polymerase mutants used in the dinucleotide expansion assay.**

T4 DNA Polymerase mutant	Location	Effect of mutation
L412M	Pol active centre	Reduced initiation of proofreading pathway
I417V	Pol active centre	Increased initiation of proofreading pathway
G255S	Exo domain	Defective in strand separation/transfer and in forming pre-exonuclease complex
W213S	Exo active centre	Reduced ability to return primer from exonuclease to polymerase active centre
D112A/E114A	Exo active centre	Cannot perform hydrolysis reaction

Another advantage in using the T4 DNA polymerase as a model is that the T4 DNA polymerase and its accessory proteins are orthologs of the human DNA polymerase delta (δ) – proliferating cell nuclear antigen (PCNA) complex (Tsurimoto and Stillman, 1990). Both protein sequence similarities (Braithwaite and Ito, 1993) and functional similarities are observed (Tsurimoto and Stillman, 1990).

Five mutant DNA polymerases were selected for this study (refer to Table I). One mutant, the D112A/E114A-DNA polymerase, is severely defective in 3'→5' exonuclease activity (Reha-Krantz and Nonay, 1993). Aspartate 112 is a ligand to divalent metal ions bound in the exonuclease active centre. In the absence of metal binding, hydrolysis activity is reduced greater than one-thousand fold. Another mutant tested is the G255S-DNA polymerase. This mutant was first isolated in a genetic selection scheme to identify mutator DNA polymerases (Reha-Krantz, 1988) and in a second selection to identify DNA polymerase active-site-switching mutants (Stocki *et al.*, 1995). The G255S-DNA polymerase is defective in moving DNA from the polymerase to the exonuclease active centre and in forming the pre-exonuclease intermediate in the proofreading pathway (Marquez and Reha-Krantz, 1996; Baker and Reha-Krantz, 1998). A third mutant is the W213S-DNA polymerase which appears to be defective in moving DNA from the exonuclease back to the polymerase active centre after the terminal nucleotide has been removed (Stocki *et al.*, 1995). Residues D112, G255, and W213 all reside in the exonuclease domain (Wang *et al.*, 1996) but residues in the polymerase active site are also predicted to affect misalignment mutagenesis. Amino acid residues L412 and I417 reside in the polymerase active centre (reviewed by Joyce and Steitz, 1995). Conservative amino acid substitutions can either increase (I417V) or decrease (L412M) exonucleolytic proofreading (Reha-Krantz *et al.*, 1993; Reha-Krantz and Nonay, 1994). If proofreading provides an opportunity for misalignment mutagenesis, then the G255S and L412M-DNA polymerases will have reduced misalignment mutagenesis since these mutants have a reduced ability to separate the DNA strands or to initiate the proofreading pathway. The W213S and I417V-DNA polymerases, on the other hand, are predicted to increase the lifetime of the strand separated complex or the opportunity to proofread.

My project was to develop an *in vitro* experimental assay and to use the mutant DNA polymerases to test the hypothesis that DNA polymerases participate in misalignment mutagenesis. The assays developed by others (for example, Kunkel, 1985 – Fig 3; Schlötterer and Tautz, 1992 – Fig 4; Petruska *et al.*, 1998) have been useful in demonstrating misalignment mutagenesis during DNA replication, but the DNA substrates used by Schlötterer and Tautz (1992) and Petruska (1998) would not be found in nature. Although dinucleotide repeats are common in natural sequences, the repeat sequences are flanked by non-repeat sequences in chromosomes. These flanking sequences are not present in the substrates used by Schlötterer and Tautz (Fig 4) and Petruska *et al.* (1998). The DNA substrate used by Kunkel (1985a&b) is a more natural substrate, but DNA sequencing is required to determine the sequence of each mutant. The DNA substrate used in this project is shown in Figure 6. This DNA substrate contains nine CA/GT repeats. The CA/GT repeats were chosen since poly $(GT)_{10\text{-}30}$ repeat sequences are one of the most abundant found in eukaryotic genomes (Hamada *et al.*, 1982). The repeat sequences are flanked by non-repeat sequences. The sequences of the template and primer strands were designed to allow specific annealing of the oligonucleotides and to eliminate the "free" ends observed with the substrates used by Schlötterer and Tautz (1992) (see Fig 4). Another feature of this DNA substrate is the single-nucleotide gap. Such a gap may arise from base excision repair (Duncan *et al.*, 1976) or may simulate DNA replication at an Okazaki fragment. Thus, more *in vivo* like conditions are mimicked by our dinucleotide expansion assay than were in previous assays.

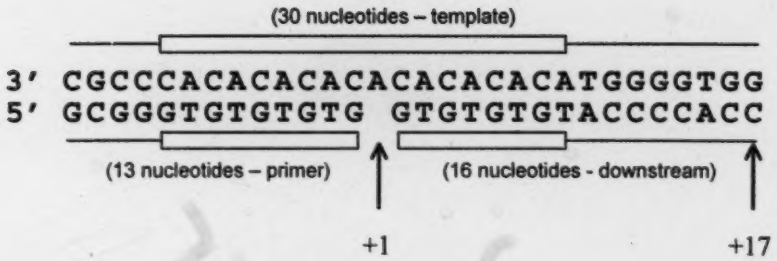


Figure 6. **DNA template-primer substrate**. The DNA substrate used in this project contains a $(CA)_9$ repeat tract, a single-nucleotide gap, and non-repetitive, 'anchored' ends. The primer is 5'[^{32}P] labeled and a downstream oligonucleotide with a 5'-phosphate is also present. Incorporation at the one-nucleotide gap is marked +1 while full-length synthesis is denoted +17. Primer extension resulting in more than nine (GT) repeats (past full-length, +17) indicates expansion mutagenesis.

EXPERIMENTAL PROCEDURES

DNA Polymerases

Purification and characterization of the wild-type bacteriophage T4 DNA polymerase and the D112A/E114A-DNA polymerase (Reha-Krantz and Nonay, 1993), the G255S-DNA polymerase (Stocki *et. al.*, 1995; Marquez and Reha-Krantz, 1996), the W213S-DNA polymerase (Stocki *et. al.*, 1995), the L412M-DNA polymerase (Reha-Krantz *et al.*, 1993), and the I417V-DNA polymerase (Reha-Krantz and Nonay, 1994) have been described.

Synthesis and Purification of DNA Oligonucleotides

Oligonucleotides were synthesized by the DNA Synthesis Lab in the Department of Biological Sciences (Table II). The oligonucleotides were purified by preparative polyacrylamide gel electrophoresis. Approximately 30 μg of DNA was loaded onto denaturing gels containing 15% polyacrylamide, 7 M urea and Tris-borate-EDTA buffer (pH 8.3). Gels were run at 55W for 4 hours. DNA bands were visualized by UV-shadowing using TLC silica gel 60 F_{254} sheets (EM Industries). Bands corresponding to full-length oligonucleotides were cut from the gel. Gel pieces were crushed and DNA was eluted with 0.3 M sodium acetate (pH 5.2) by shaking at 37 °C overnight. DNA was purified through Sep-Pak cartridges (Millipore). Briefly, the column was charged with 100% methanol and equilibrated with buffer (100 mM Tris [pH 8.0] and 1 mM EDTA). DNA was loaded onto the column slowly, and washed with buffer (100 mM Tris [pH 8.0] and 1 mM EDTA) followed by H_2O. DNA was eluted from the column with 40% methanol. The first three fractions (1 mL each) were collected and dried in a speed vacuum dryer. The purified oligonucleotides were dissolved in 50 uL H_2O and stored at –20 °C. DNA concentrations were determined by spectrophotometry at 260 nm. Extinction coefficients of the oligonucleotides were calculated based on nucleotide composition.

Table II. DNA Oligonucleotides Synthesized for the Dinucleotide Expansion Assay

Expansion assay	
LRK28 **(30-mer – template)**	3' CGCCCACACACACACACACACATGGGGTGG
LRK27 **(13-mer – primer)**	5' GCGGGTGTGTGTG
LRK103-2 **(16-mer – downstream)**	5' ‡GTGTGTGTACCCCACC
Phosphorothioate DNA	
LRK135 **(30-mer – template)**	3' C•GCCCACACACACACACACACATGGGGTGG
LRK156 **(13-mer – primer)**	5' GCGGGTGTGTGT•G
LRK139 **(16-mer – downstream)**	5' ‡GTGTGTGTACCCCAC•C
DNA Markers	
LRK116 **(30-mer)**	5' GCGGGTGTGTGTGTGTGTGTGTACCCCACC
LRK118 **(32-mer)**	5' GCGGGTGTGTGTGTGTGTGTGTGTACCCCACC
LRK120 **(34-mer)**	5'GCGGGTGTGTGTGTGTGTGTGTGTGTACCCCACC
LRK117 **(30-mer)**	5' GCGGGTGTGTGTGTGTGTGTGTGTGTGTGT
LRK119 **(32-mer)**	5' GCGGGTGTGTGTGTGTGTGTGGTGTGTGTGT
LRK121 **(34-mer)**	5' GCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT

‡ - phosphate
• - phosphorothioate

Removal of the R_p Isomer from Phosphorothioate Oligonucleotides

Oligonucleotides synthesized with phosphorothioates contain both R_p and S_p isomers due to the chirality of the synthesis reaction. Oligonucleotides containing the R_p isomer are degraded by the 3'→5' exonuclease of T4 DNA polymerase while the S_p isomer is resistant (Eckstein, 1985; Spitzer and Eckstein, 1988). Exonuclease reactions contained 0.5 mM dithiothretol, 67 mM Tris (pH 8.0), 16.7 mM $(NH_4)_2SO_4$, 200 μg/mL bovine serum albumin, 6.7 mM $MgCl_2$, and a 5:1 ratio of wild-type T4 DNA polymerase to DNA oligonucleotide. Reactions were incubated at 37 °C for 1.5 hours. The remaining S_p isomer oligonucleotides were purified by polyacrylamide gel electrophoresis as described above.

End Labeling and Annealing Reactions

Oligonucleotides, 7.5 picomoles, were [^{32}P] labeled at the 5'-end by T4 DNA polynucleotide kinase (Boehringer Manheim) with 9 pmol [γ^{32}P]ATP (3000 Ci/mmol, Dupont NEN) (Maniatis, 1982). Kinased oligonucleotides were annealed with 1.5-fold excess template to primer in 67 mM Tris-HCl buffer (pH 8.0). The solution was heated to 70 °C and cooled slowly to room temperature.

Dinucleotide Expansion Assay

Labeled primer-template substrate [3.75 nM] (see Table II and Fig 6) was pre-incubated at 35°C for 1 min in a 100 μL reaction mixture containing 100 μM dNTPs, 0.5 mM dithiothretol, 67 mM Tris (pH 8.0), 16.7 mM $(NH_4)_2SO_4$, 200 μg/mL bovine serum albumin, 6.7 mM $MgCl_2$. T4 DNA polymerase, at a final concentration of 7.5 nM, was added to initiate replication. Reactions were incubated at 35°C and stopped by mixing 8 μL of reaction mix with 8 μL loading dye (10% formamide, 0.1% cyanol blue, 0.1% xylene blue). Expansion products were analyzed by gel electrophoresis on denaturing 15% polyacrylamide gels containing 7 M urea, and Tris-borate-EDTA buffer (pH 8.3) or similar gels which also contained 25% formamide. Gels were run at 45 W for 75 min, dried, and exposed to phosphorimager screens (Molecular Dynamics).

Maxam-Gilbert Sequencing of Expansion Products

Selected 5'[^{32}P] labeled dinucleotide expansion products were purified by gel electrophoresis on 15% denaturing acrylamide gels (7 M urea and Tris-borate-EDTA buffer [pH 8.3]) containing 25% formamide. The gels were frozen and exposed overnight to X-ray film (Kodak-XR), to visualize the bands. The desired band was cut from the gel. DNA from the band was eluted by crushing the gel slices and shaking the slices in 0.3 M sodium acetate (pH 5.2) overnight. DNA was purified as described before (see Synthesis and Purification of DNA Oligonucleotides). Samples having a total minimum of 300 cpm were deemed suitable for sequencing.

Maxam and Gilbert (1980) sequencing reactions were performed as described (Current Protocols in Molecular Biology, 1996) except that the T reaction was done with $KMnO_4$ as described by Rubin and Schmid (1980). Briefly, G was modified by the addition of dimethyl sulfate [DMS] (Sigma-Aldrich), C was modified by the addition of hydrazine [HZ] (Sigma-Aldrich), and both A and G were modified by the addition of formic acid (Sigma-Aldrich). The G and T reactions were stopped by addition of 50 µL DMS Stop buffer (1.5 M sodium acetate [pH 7.0], 1.0 M β-mercaptoethanol, 100 µg/mL tRNA) and 500 µL of ice cold 100% ethanol. The A+G and C reactions were stopped by addition of 200 µL HZ Stop buffer (0.3 M sodium acetate [pH 7.0], 0.1 mM EDTA, 25 µg/mL tRNA) and 750 µL ice cold 100% ethanol. Stopped reactions were placed at –80 °C for 45 min followed by centrifugation at 13,000 rpm for 45 min at 4 °C. After removal of the supernatant and drying, 100 µL of 10% piperidine (Sigma-Aldrich) was added to each tube and incubated at 90 °C for 30 min. Piperidine was removed by evaporation and sequencing products were washed once with 40 µL and twice with 30 µL millipore H_2O. 'Sequencing' products were analyzed by gel electrophoresis on 15% denaturing polyacrylamide gels (7 M urea and Tris-borate-EDTA buffer [pH 8.3]) containing 25% formamide, dried, and exposed to phosphorimager screens (Molecular Dynamics).

RESULTS

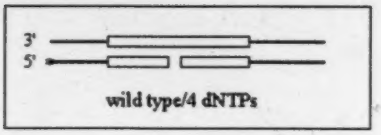
In vitro dinucleotide expansion assay

The DNA substrate used in this project (Fig 6) was designed to approximate a chromosomally located dinucleotide repeat sequence, unlike those used in previous assays (Schlötterer and Tautz, 1992; Petruska *et al.*, 1998). Nine CA/GT repeats are flanked by non-repeat DNA and a single nucleotide gap is present in the substrate. The primer-terminus is located within the repeats. Elongation of the primer-terminus is predicted to extend the primer by one nucleotide to fill in the gap (+1). Further elongation requires displacement and replication through the downstream oligonucleotide. Products were separated and visualized by polyacrylamide gel electrophoresis on 15% acrylamide-7 M urea gels. Filling of the gap (+1) occurred (Fig 7). However, full-length synthesis (+17) and incorporation of nucleotides past that templated by the DNA substrate (full length, +17) was also observed within 15 s (Fig 7). Further expansion occurred as the time course progressed, and four to five extra dinucleotide repeats were added after 120 min (Fig 7). The proposal that dinucleotide expansion, due to misalignment mutagenesis, was responsible for extension beyond full-length (+17) is based on the fact that longer products appear to be extended in units of two. A pattern of a light band followed by a dark band is clearly seen. Dark bands are thought to correspond to an additional GT repeat followed by complete synthesis of the non-repeat downstream sequence. The light bands also contain an additional GT repeat; however, replication of the downstream non-repeat sequence is incomplete in these bands. This occurs occasionally if the DNA polymerase becomes unstable and dissociates from the DNA when it reaches the end of the template. Thus, light bands correspond to products that are one nucleotide short of the complete sequence.

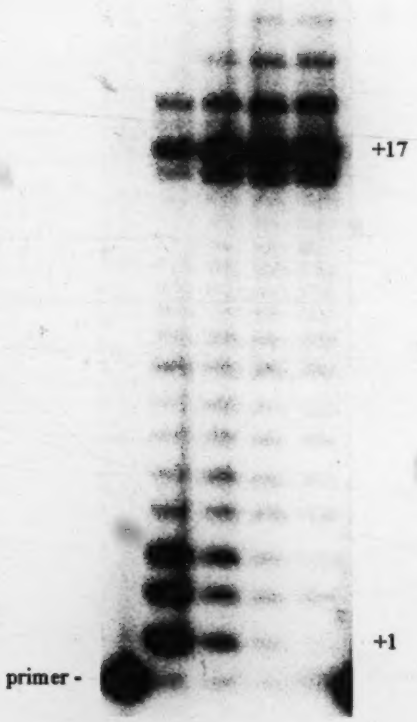
Sequence of expansion products

While it was presumed that additional dinucleotide repeats were present in the expansion (longer than full-length) products, the exact composition of these products was unknown at this point. To determine the nature of the products longer than the template length, synthesized oligonucleotides of known composition and size were

Figure 7. **Dinucleotide expansion catalyzed by wild type T4 DNA polymerase**. The primer oligonucleotide of the single-gapped DNA substrate was 5'-end labeled with [^{32}P]. Wild type DNA polymerase was incubated with the DNA substrate and all four dNTPs. Reactions were incubated for 15 s, 10 min, 30 min, and 120 min. Products were resolved by gel electrophoresis. The 3'-end of the primer-terminus is labeled "primer". DNA replication to the end of the template is denoted +17 (full-length). Products observed past full-length result from expansion of the (GT) repeat tract. Detailed reaction conditions are described under "Experimental Procedures".



15" 10' 30' 120'



run next to the expansion products on polyacrylamide gels. The oligonucleotide markers consisted of either the primer non-repeat sequence followed by GT repeats or DNA containing both upstream and downstream non-repeat sequences with additional GT repeats internal to the non-repeat sequences (see Table II). Expansion products appeared to align with the markers containing additional internal GT repeats and the downstream non-repeat sequence (Fig 8). The Maxam and Gilbert (1980) sequencing method was subsequently used to ascertain the exact nucleotide composition of the expansion products (Fig 9). Sequencing of the purified expansion product confirmed the initial observations that the extra dinucleotide GT repeat was internal to the non-repeat sequences. The synthesized DNA substrate contains a $(GT)_9$ tract. A non-interrupted stretch of $(GT)_{10}$ was seen in Figure 9 indicating the incorporation of an extra dinucleotide repeat.

Role of the downstream oligonucleotide and DNA polymerase proofreading

To determine if the downstream oligonucleotide was required for dinucleotide expansion, a 30-mer, 5'[^{32}P] labeled primer-template without the 16-mer downstream oligonucleotide was prepared (see Table II). Under reaction conditions (as described in 'Experimental Procedures') which included wild type T4 DNA polymerase and all four deoxynucleotides, replication of the complete template (full-length, +17 position) occurred (Fig 10A). No dinucleotide expansion was observed. Therefore, a downstream oligonucleotide is needed for expansion to occur.

The need for slowed translocation of the polymerase was explored by determining the importance of DNA polymerase exonuclease proficiency on the gapped DNA substrate. This was addressed by substituting T4 exonuclease-deficient D112A/E114A-DNA polymerase for wild type polymerase in an identical assay. The difference between these two polymerases is their ability to proofread (see 'Introduction'). When supplied with a gapped substrate and all four nucleotides, the exonuclease-deficient D112A/E114A-DNA polymerase was unable to catalyze dinucleotide expansion (Fig 10B). Although replication was observed past those templated by the DNA substrate (full-length, +17), the +18 product is likely the incorporation of a terminal non-templated nucleotide. Such blunt end extension has

Figure 8. **Determination of expansion product composition**. Wild type T4 DNA polymerase was incubated in reactions with the single-gapped DNA substrate and all four dNTPs for 180 min. Reaction products were run with synthesized oligonucleotides of known size and composition (see Table II) on polyacrylamide gels. The 'marker' oligonucleotides consisted of either pure GT repeats originating with the primer 'anchor' sequence or GT repeat sequences with both primer and downstream 'anchor' sequences. Expansion products appear to align with the synthetic DNA 'markers' containing both primer and downstream 'anchor' sequences.

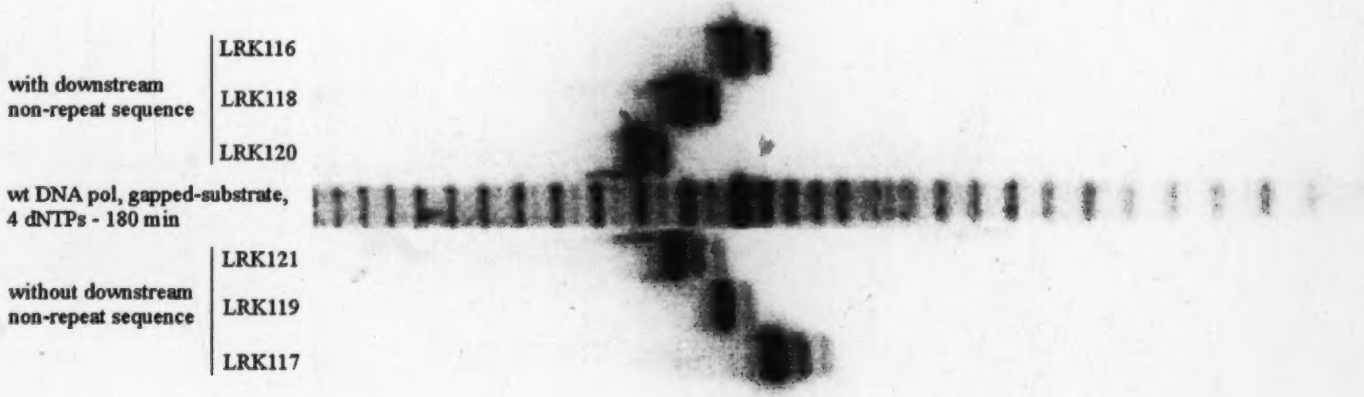
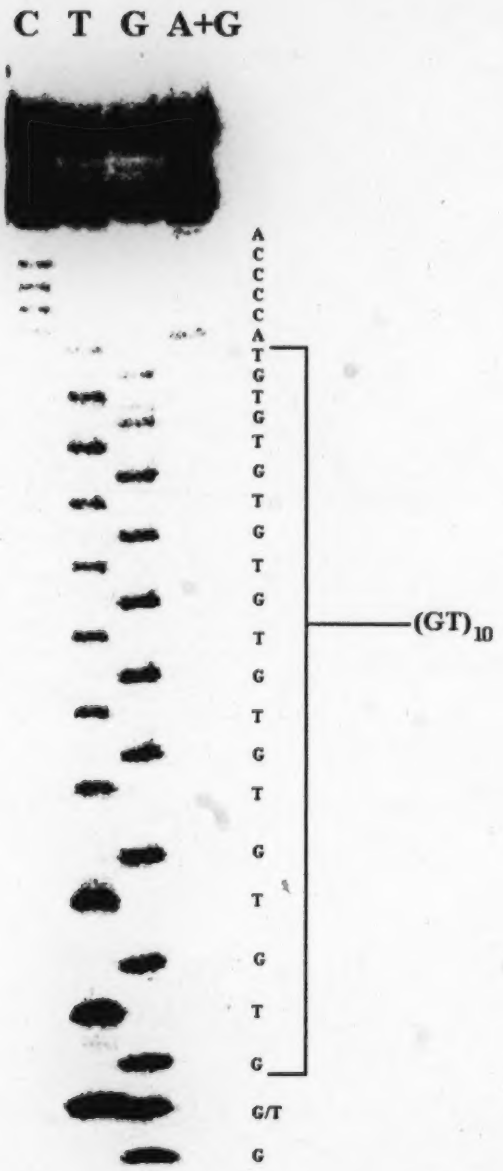

LRK116
with downstream non-repeat sequence
LRK118
LRK120
wt DNA pol, gapped-substrate, 4 dNTPs - 180 min
LRK121
without downstream non-repeat sequence
LRK119
LRK117

Figure 9. **Determination of expansion product sequence**. The Maxam and Gilbert chemical sequencing method (1980) was used to determine the sequence of an expansion product containing one extra dinucleotide repeat. Expansion products from reactions containing wild type T4 DNA polymerase, the single-gapped DNA sequence and all four dNTPs were resolved by gel electrophoresis. Expansion product corresponding to +19 (one extra GT repeat) was cut from the gel and purified. The 'pure' expansion product was then sequenced as described under "Experimental Procedures". The large band at the 5'-end (bottom) of the sequence is labeled G/T due to incorrect synthesis of the template oligonucleotide. The template was synthesized with an adenine residue in addition to the cytosine residue desired at this position.


C T G A+G
A C C C C A T G T G T G

been detected for the exonuclease-deficient Klenow fragment (Clark *et al.*, 1987) and *Thermus aquaticus* (Taq) DNA polymerase (Clark, 1988).

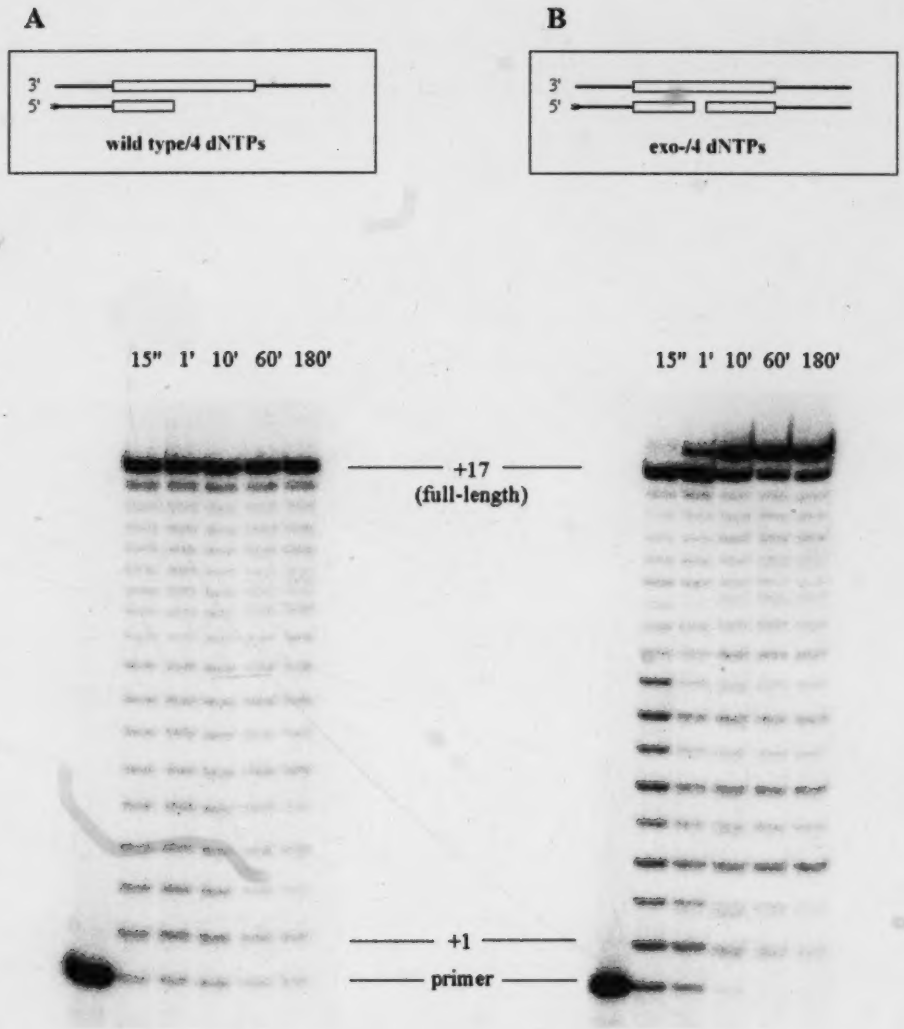
Phosphorothioates to block 3'→5' exonuclease activity

The DNA primer-template used contains non-repeat sequences designed to allow precise annealing of the primer and template strands. However, the potent 3'→5' exonuclease activity of T4 DNA polymerase can remove the oligonucleotide 3'-ends. Degradation of the both template and downstream non-repeat sequence can create a primer-template substrate consisting exclusively of repeats, and "free" ends, like that described by Schlötterer and Tautz (1992) (Fig 4).

To negate the effect of the 3'→5' exonuclease activity, the template and downstream oligonucleotides were synthesized with phosphorothioates at various positions along the oligonucleotides (see Fig 11 and Table II). Phosphorothioates replace one of the non-esterified oxygen atoms normally present in the phosphodiester backbone with a sulfur atom (reviewed by Frey, 1989). The stereochemistry of this reaction creates two isomers denoted R_p and S_p. The R_p isomer is susceptible to degradation by the 3'→5' exonuclease, but the S_p isomer is resistant. In fact, DNA polymerase is not only unable to degrade the DNA, but it is also unable to bind the S_p isomer in the exonuclease active centre (Brautigam and Steitz, 1998).

Placement of a phosphorothioate at the terminal 3'-end of either the template, downstream, or primer oligonucleotide in the single-gapped DNA substrate did not appear to affect the amount of dinucleotide expansion seen (Fig 12A, B, C) when compared to wild type DNA polymerase and the unmodified DNA substrate under similar reaction conditions (Fig 7). The kinetics of the reactions with the DNA substrate containing the phosphorothioate modified template (Fig 12A) were slightly slower than that observed for the two other modified substrates (Fig 12B & C). Less full-length synthesis (+17) was seen within the first 15 s and only one additional GT repeat was added by 10 min (Fig 12A). However, the characteristic pattern of dinucleotide expansion products was present. Therefore, the DNA substrate does not

Figure 10. **Determination of the importance of the downstream oligonucleotide and DNA polymerase proofreading for dinucleotide expansion**. Panel A. Wild type T4 DNA polymerase was incubated with a single stranded DNA substrate and all four dNTPs. Panel B. A second reaction contained exonuclease-deficient D112A/E114A-DNA polymerase, the single-gapped DNA substrate and all four dNTPs. Products of both reactions were separated by gel electrophoresis. The end of the primer-terminus is labeled 'primer'. For both reactions, incorporation of the first nucleotide is denoted +1 while synthesis of the complete template is denoted +17 (full-length).


A
3'
5'
wild type/4 dNTPs
B
3'
5'
exo-/4 dNTPs
15" 1' 10' 60' 180'
15" 1' 10' 60' 180'
+17
(full-length)
+1
primer

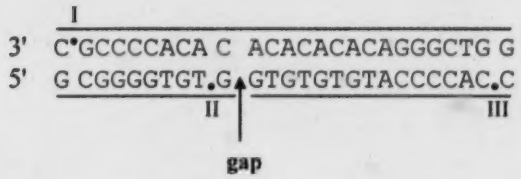
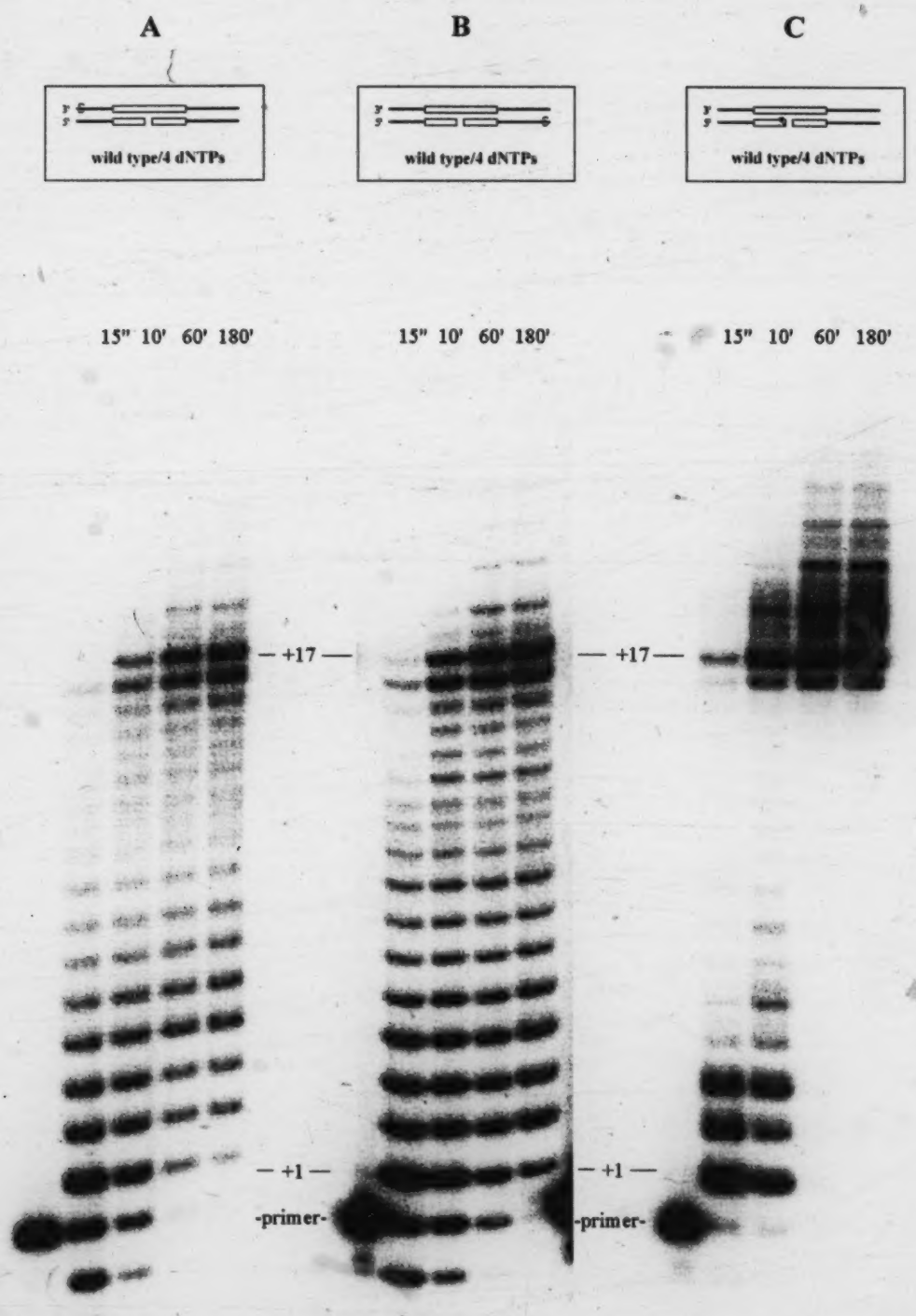


Figure 11. **Placement of phosphorothioates along the DNA substrate**. DNA substrates were synthesized with phosphorothioates at the 3'-ends of the template (I), downstream (III) and primer (II) oligonucleotides to block 3'→5' exonuclease activity. The positions of the phosphorothioates are indicated as dots with corresponding roman numerals.

Figure 12. **Using phosphorothioates to determine the importance of 3'→5' exonucleolytic proofreading**. Phosphorothioates were placed at the 3'-most linkage of the template (panel A), downstream (panel B), or primer (panel C) oligonucleotide. Reactions contained wild type T4 DNA polymerase, the single-gapped substrate with modification described above, and all four dNTPs. Reactions were incubated for 15 s, 10 min, 60 min, and 180 min and the products were resolved by gel electrophoresis. The end of the primer-terminus is labeled 'primer'. Filling of the gap is denoted +1. Replication to the end of the template is denoted +17 (full-length) while synthesis past this is indicative of dinucleotide expansion.


A
B
C
wild type/4 dNTPs
wild type/4 dNTPs
wild type/4 dNTPs
15" 10' 60' 180'
15" 10' 60' 180'
15" 10' 60' 180'
+17
+1
-primer-
+17
+1
-primer-

require trimming of the non-repeat sequences at either template or downstream oligonucleotide end for dinucleotide expansion to occur.

Importance of Strand Separation

The expansion assay was also performed with the mutant T4 DNA polymerases described previously (Table I), the single-gapped DNA substrate and all four dNTPs. The L412M, I417V, G255S, W213S and D112A/E114A-DNA polymerases all affect some aspect of the proofreading pathway. The preliminary results are shown in Figure 13.

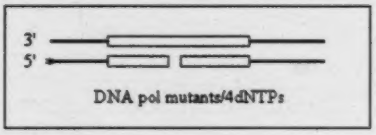
Dinucleotide expansion was observed, as previously seen (Fig 7), for wild type T4 DNA polymerase (Fig 13). Filling of the single-nucleotide gap (+1) and replication to full-length (+17) was observed in the first 15 s. Polymerase "pausing" occurred after filling of the gap (Fig 13, 15 s). An extra GT repeat was also added within 15 s. As time progressed, more primer was extended to full-length (+17) and extra GT repeats corresponding to dinucleotide expansion were created (Fig 13, 30 min & 120 min).

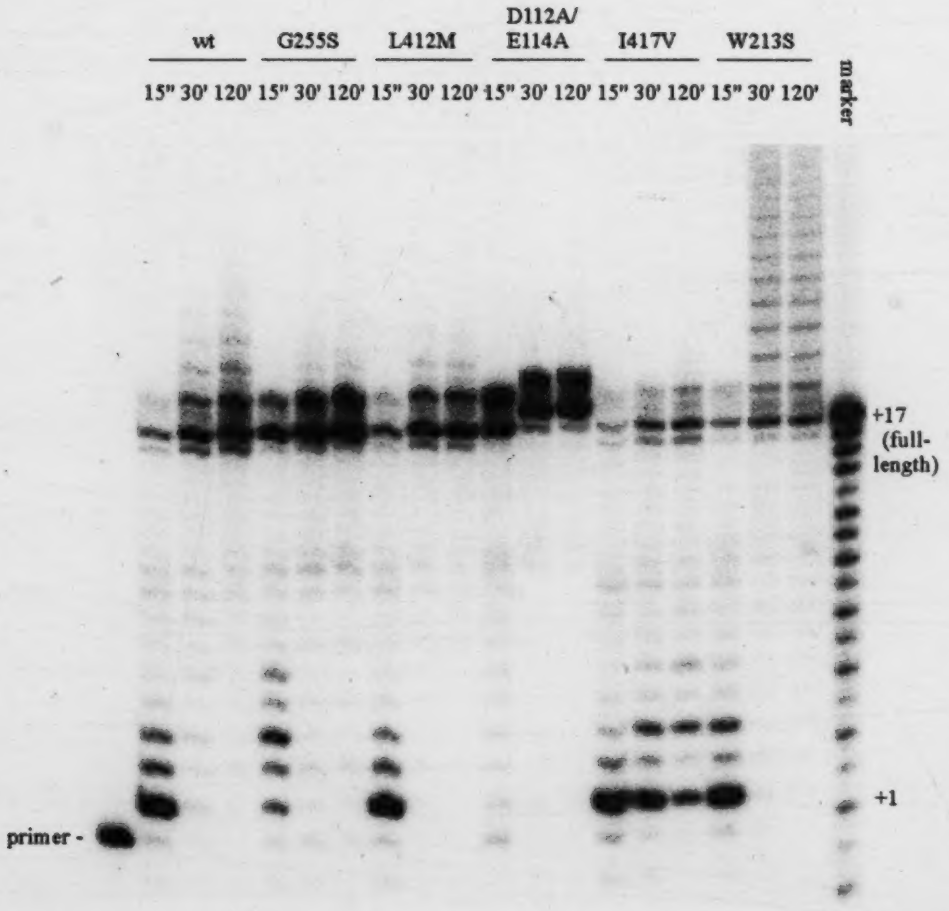
The G255S-DNA polymerase mutant is predicted to be unable to promote expansion mutagenesis. Limited expansion was observed with the gapped DNA substrate (Fig 13). Replication to full-length (+17) was observed within 15 s with the addition of an extra dinucleotide repeat. More "pausing" by the polymerase was seen through the first four nucleotides past the gap (Fig 13, 15 s). After 30 min of incubation, two additional dinucleotide repeats were visualized (Fig 13). The number of additional repeats did not increase after 120 min (Fig 13).

L412M-DNA polymerase is predicted to promote less dinucleotide expansion due to reduced exonucleolytic proofreading. Replication filled in the gap (+1) with extension to full-length (+17) and addition of an extra dinucleotide repeat (Fig 13). DNA polymerase "pausing" was again observed within the first few nucleotides (Fig 14, 15 s). As time progressed, more primer was extended to full-length (+17) product and additional expansion products were created (Fig 13, 30 min & 120 min).

The D112A/E114A-DNA polymerase mutant previously did not catalyze dinucleotide expansion (Fig 10B). In the current reaction this DNA polymerase

Figure 13. **Comparison of T4 DNA polymerase mutants' ability to promote dinucleotide expansion**. Wild type and mutant T4 DNA polymerases G255S, L412M, D112A/E114A, L417V, and W213S were incubated in separate reactions with the single-gapped DNA substrate and all four dNTPs. Incubation were run for 15 s, 30 min, and 120 min. Products were resolved by gel electrophoresis. The end of the primer-terminus is labeled 'primer'. Full-length synthesis was determined by running a synthesized oligonucleotide, corresponding to full-length, alongside the reaction products and is denoted +17.


3'
5'
DNA pol mutants/4dNTPs


wt
G255S
L412M
D112A/
E114A
I417V
W213S
15" 30' 120' 15" 30' 120' 15" 30' 120' 15" 30' 120' 15" 30' 120' 15" 30' 120'
marker
+17
(full-
length)
+1
primer -

synthesized full-length (+17) product and added one extra dinucleotide repeat within 15 s (Fig 13). As the incubation time progressed, the characteristic addition of a non-templated nucleotide occurred, due to reasons previously explained. The extra dinucleotide repeat had an additional non-templated nucleotide added after 30 min. Little DNA polymerase "pausing" was observed. The discrepancy (addition or lack of a dinucleotide repeat) between the two D112A/E114A-DNA polymerase results (Fig 10B & 13) is addressed in the 'Discussion'.

I417V-DNA polymerase is predicted to catalyze dinucleotide expansion. Filling of the gap (+1) and polymerase "pausing" appeared as a major component of this reaction (Fig 13). Full-length synthesis (+17) was observed with some addition of an extra dinucleotide repeat within 15 s. At 120 min, a very small amount of a second dinucleotide repeat was observed (Fig 13).

W213S-DNA polymerase is predicted to increase the amount of dinucleotide expansion observed. In the first 15 s, the gap was primarily filled (+1) with some "pausing" of the DNA polymerase (Fig 13). Full-length synthesis (+17) and an extra dinucleotide repeat was also seen. At 30 min and 120 min, substantially more expansion is observed than was seen for wild type DNA polymerase, following the prediction that W213S-DNA polymerase should catalyze more dinucleotide expansion (Fig 13).

DISCUSSION

Characterization of the *in vitro* dinucleotide expansion assay

An *in vitro* dinucleotide expansion assay was developed that has significant improvements compared to previous assays. The assay began with the construction of a single-gapped DNA substrate designed to allow for specific annealing of the DNA oligonucleotides and to mimic *in vivo* conditions of chromosome replication involving Okazaki fragments. These two conditions were fulfilled by synthesizing a DNA substrate with a repeat sequence consisting of $(GT)_9$, flanked on both sides by non-repetitive, GC rich sequences and a one-nucleotide gap in the region of the GT repeats (Fig 6). Similar assays have previously been conducted (Schlötterer and Tautz, 1992 – Fig 4; Petruska *et al.*, 1998), however, in both instances the DNA substrates used did not resemble structures that would be found in a chromosomal location.

Dinucleotide expansion was first observed using our *in vitro* primer-extension assay, in the presence of wild type T4 DNA polymerase, the single-gapped DNA substrate described above, and all four deoxynucleotides (Fig 7). This result was surprising for two reasons. First, synthesis to the end of the DNA template by DNA polymerase requires displacement of the downstream oligonucleotide. In the absence of DNA helicase, T4 DNA polymerase has not previously been observed to displace more than the first few nucleotides of a downstream oligonucleotide (Dong *et al.*, 1996). Secondly, in the presence of all four dNTPs, T4 DNA polymerase should faithfully replicate the DNA substrate up to the end of template (full-length, +17) only. We were surprised to see that wild type polymerase had incorporated an additional dinucleotide repeat within 15 s and more than five dinucleotides by 120 min (Fig 7). The chemical sequencing technique (Maxam and Gilbert, 1980) was used to determine that the additional nucleotides were G and T, located in the original GT repeat tract, followed by the downstream non-repeat sequence (Fig 9).

The significance of DNA polymerase exonucleolytic proofreading was subsequently observed using exonuclease-deficient D112A/E114A-DNA polymerase. In the presence of all four deoxynucleotides and the single-gapped DNA substrate, D112A/E114A-DNA polymerase was unable to promote dinucleotide expansion (Fig

10B). The difference with these results and that observed in Figure 13 is discussed below with the T4 DNA polymerase mutants. Little "pausing" (formation of intermediate products) occurred with the exonuclease-deficient DNA polymerase (Fig 10B) while substantially more intermediate products formed when wild type DNA polymerase replicated the same substrate (Fig 7 & 13). Since the mutant D112A/E11A-DNA polymerase is 3'→5' exonuclease deficient, it is unable to proofread either correct or incorrect nucleotides that have been incorporated during replication. With a decrease in the opportunity to proofread, the exonuclease-deficient D112A/E114A-DNA polymerase is not able to undergo the "idling" reaction described previously. Exonuclease-deficient D112A/E114A-DNA polymerase appears to translocate through the downstream oligonucleotide more easily and full-length synthesis out competes expansion. Wild type polymerase, on the other hand, can proofread at any time during DNA replication, even in the absence of misincorporated nucleotides. DNA polymerases also have difficulty replicating sequences consisting of repeats and thus more DNA polymerase "pausing" is observed. The increased polymerase "pausing" within the repeat sequence may allow more proofreading to occur, allowing for strand separation and misalignment, and expansion by wild type DNA polymerase.

Another important aspect of this reaction is the presence of the downstream oligonucleotide. In the absence of the downstream oligonucleotide, wild type T4 DNA polymerase can no longer catalyze dinucleotide expansion (Fig 10A). Presumably, translocation by the wild type DNA polymerase is less hindered in the absence of a downstream oligonucleotide and the competition between full-length and expansion synthesis is shifted toward increased full-length synthesis. The downstream oligonucleotide forces more DNA polymerase "pausing" and thus more opportunity for an "idling" reaction to occur. "Idling" by the DNA polymerase, and the ability to proofread are needed in order to provide enough opportunity for the DNA polymerase to promote dinucleotide expansion. A possible scenario for strand misalignment of CA/GT repeats, promoted by DNA polymerase proofreading, is illustrated in Figure 14.

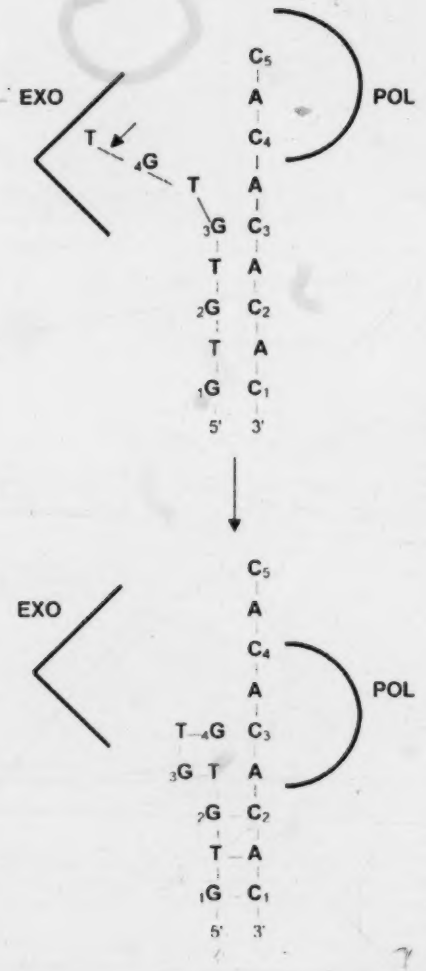


Figure 14. **Model for the strand separation and misalignment of CA/GT repeats during proofreading.** During proofreading, the primer-terminus is separated from the template strand by three or four nucleotides and moved from the polymerase to exonuclease active centre. Excision of the terminal nucleotide (T) occurs. The DNA polymerase may diffuse or slide backwards, forcing a misalignment (extrahelical loop) when the primer strand is returned to the polymerase active centre.

Exonucleolytic proofreading at all of the 3'-ends was also investigated. Since T4 DNA polymerase contains a potent 3'→5' exonuclease activity, the 3'-ends of the DNA substrate are susceptible to degradation. Removal of the non-repeat sequences can allow for non-specific annealing, creating "free"' ends such as those present in the Schlötterer and Tautz (1992) assay (Fig 4). The need for exonucleolytic trimming of the DNA substrate in the current assay was determined by placing phosphorothioates at all of the 3'-ends in our DNA substrate. The presence of the S_p isomer of this sulfur moeity in the phosphodiester backbone is an inhibitor of 3'→5' exonuclease activity (Eckstein, 1985; Spitzer and Eckstein, 1988). Placement of a phosphorothiate at the 3'-end of either the template, downstream, or primer oligonucleotide did not eliminate dinucleotide expansion catalyzed by wild type DNA polymerase (Figures 12A, B, C). Therefore, DNA strand misalignment is possible in the presence of the non-repeat sequences and exonucleolytic trimming of the template 3'-end is not required for dinucleotide expansion to occur, but appear to assist the extent of the reaction.

Probing the importance of proofreading with T4 DNA polymerase mutants

After the critical aspects of the *in vitro* assay were determined, the hypothesis that misalignment mutagenesis is assisted by the DNA polymerase during exonucleolytic proofreading was tested. A comparison of functionally distinct mutant T4 DNA polymerases was employed to probe the mechanism involved in misalignment mutagenesis. Although DNA replication has been linked with misalignment mutagenesis (Streisinger, 1966; Kunkel, 1985; Schlötterer and Tautz, 1992; Strand *et al.*, 1993) a mechanism by which misalignment mutagenesis occurs has not been elucidated. DNA polymerases may play a part in catalyzing strand misalignment or, as proposed by others (Schlötterer and Tautz, 1992; Petruska *et al.*, 1998), the DNA strand misalignment simply occurs spontaneously, following the dissociation of the DNA polymerase. Our hypothesis that DNA polymerases play an active role in catalyzing strand misalignment arose because DNA polymerases with 3'→5' exonuclease actively assist strand separation during exonucleolytic proofreading. In order to remove a misincorporated nucleotide, the 3'-end of the

primer strand must be separated from the template strand and transferred from the polymerase to exonuclease active centre. After the excision of the incorrect nucleotide from the primer-terminus, the strand is transferred back to the polymerase active centre (see Fig 5). Thus, strand separation during exonucleolytic proofreading increases the opportunity for strand misalignment. Five mutant DNA polymerases that affect the proofreading pathway were tested using the *in vitro* dinucleotide expansion assay and the preliminary results are discussed below.

The G255S-DNA polymerase mutant encodes a substitution for a residue that resides in the loop structure found between the spatially distinct polymerase and exonuclease domains in the DNA polymerase enzyme (Wang *et al.*, 1996). This loop structure is thought to be involved in the separation and switching of the DNA primer strand to the exonuclease centre for proofreading (Marquez and Reha-Krantz, 1996). The serine substitution is proposed to destabilize the loop structure. If strand separation is responsible for promoting repeat expansion, this mutant DNA polymerase is predicted to be unable to promote dinucleotide expansion. Interestingly, an extra dinucleotide repeat was observed within 15 s and a few additional repeats were produced by 30 min incubation (Fig 13). Thus, the results observed for the G255S-DNA polymerase do not follow our original hypothesis, however a modified model for dinucleotide expansion is explained below.

L412M-DNA was initially discovered as a second site suppressor of excessive exonucleolytic proofreading. This mutation is found in the polymerase active centre, causing the balance in activity to be shifted from exonucleolytic proofreading to polymerization (Reha-Krantz *et al.*, 1993). A reduction in proofreading results in less opportunity to separate the DNA strands and based on the original hypothesis, less opportunity to catalyze dinucleotide expansion. Although the L412M-DNA polymerase still catalyzed dinucleotide expansion, less expansion products were produced than observed for wild type DNA polymerase (Fig 13) under the same reaction conditions.

The D112A/E114A-DNA polymerase mutant was described previously. Although this mutant is unable to catalyze the hydrolysis component of the proofreading pathway, it is able to separate the DNA strands and switch the primer-

terminus from the polymerase to exonuclease active site. This exonuclease-deficient DNA polymerase was previously unable to catalyze any dinucleotide expansion (Fig 10B), contrary to initial predictions. However more recent experiments demonstrate the addition of one extra dinucleotide repeat (Fig 13). This inconsistency in observations is likely due to differences in reaction temperature or enzyme concentration, although these possibilities have not yet been rigorously tested. If D112A/E114A-DNA polymerase does in fact catalyze dinucleotide expansion, this result suggests that there are in fact two steps in our dinucleotide expansion model. The first step is an initial expansion of only one dinucleotide repeat. This single addition occurs quickly, within the first 15 s. The second step occurs later, and involves the larger expansion products observed, for example, those at 30 min with W213S-DNA polymerase. This second step may require trimming at the 3'-ends so that more strand slippage is possible.

Another amino acid substitution in the polymerase active centre, I417V, destabilizes interactions with the primer-terminus within this active centre (Reha-Krantz and Nonay, 1994). This mutation increases the opportunity for the enzyme to form complexes in which the primer-terminus resides in the exonuclease active centre. This enzyme should therefore be more proficient in catalyzing dinucleotide expansion. In the reaction containing I417V-DNA polymerase, the single nucleotide gap is filled in (+1), however, very little of the +1 product is extended further to full-length (+17), even after 120 min (Fig 13). This is indicative of a strong "idling" reaction, due to excessive proofreading. A very small amount of expansion products are present after 120 min. If only a few extra repeats are added, the excessive 3'→5' exonuclease activity of I417V-DNA polymerase presumably degrades these products.

W213S-DNA polymerase is deficient in returning the primer-terminus from the exonuclease active centre back to the polymerase active centre after the terminal nucleotide has been removed. This mutant polymerase increases the time during which the DNA strands are separated and therefore should increase the opportunity for strand misalignment. Increased dinucleotide expansion is predicted and was observed (Fig 13). By 30 min, W213S-DNA polymerase had added ten additional dinucleotide repeats, whereas wild type DNA polymerase had only incorporated four

or five extra dinucleotide repeats by this time. W213S-DNA polymerase also catalyzed the largest expansion products (Fig 13).

Three of the mutants tested, L412M-DNA polymerase, I417V-DNA polymerase and G255S-DNA polymerase did not follow our initial predictions. I would like to suggest that the discrepancies observed are likely due to an over simplified identification of the mutants. For example, I417V-DNA polymerase is capable of conducting all of the steps involved in proofreading and is more likely to proofread than wild type DNA polymerase. However, almost no dinucleotide expansion was observed in assays conducted with this mutant (Fig 13). An important factor, neglected in the original prediction was the increased ability of this mutant to degrade DNA due to an increase in proofreading activity. Therefore, if this DNA polymerase was able to add extra dinucleotide repeats they would most likely be removed by the 3' $\rightarrow$ 5' exonuclease activity and thus this mutant is inaccurately described as being unable to promote dinucleotide expansion. On the other hand, the G255S-DNA polymerase mutant is deficient in the separation of duplex DNA and thus predicted to be unable to promote dinucleotide expansion. However, this prediction was not observed (Fig 13) as well. The G255S-DNA polymerase mutant is able to directly bind DNA into the exonuclease active centre if the DNA has been partially separated due to the "breathing" that occurs at the ends of DNA duplexes. This could account for the small amount of dinucleotide expansion observed with this mutant. It is important to note that these other factors exist and that they may contribute to the amount of dinucleotide expansion observed for each mutant. Thus, the initial predictions based exclusively on the ability of the mutants to perform the strand separation step of the proofreading pathway need to be modified to take these additional factors into consideration.

Summary and Future Directions

A greatly improved *in vitro* assay that demonstrates DNA polymerase catalyzed dinucleotide expansion was developed. The major improvement in the assay is in the design of the DNA substrate (Fig 6). The downstream oligonucleotide is extremely important since it forces increased polymerase "idling". This increased

"idling" by the DNA polymerase promotes more proofreading and thus an increased opportunity for strand separation and misalignment. Finally, exonucleolytic trimming at the template or downstream 3'-end is not required for expansion to occur.

The use of functionally distinct T4 DNA polymerase proofreading mutants (Table I) allowed for a preliminary comparison of their ability to catalyze dinucleotide expansion. The results for the T4 DNA polymerase mutants are different from those of wild type DNA polymerase. Therefore, the mutants can be used to probe the mechanism of dinucleotide expansion. It is important to note that the mutants can change more than one parameter of DNA polymerase function. These factors all affect the ability of the DNA polymerase to catalyze dinucleotide expansion and further work is required to discern them. However, the differences observed between the mutants indicate a role for DNA polymerase in strand misalignment and expansion mutagenesis.

To further investigate the importance of proofreading for catalyzing expansion mutagenesis and to see the subtle differences among the mutant DNA polymerases a different substrate that would allow us to better isolate the reaction at the primer-terminus and eliminate DNA polymerase binding and activity at additional 3'-ends could be constructed. A circular DNA substrate, containing a CA/GT repeat sequence and a single-nucleotide gap within the repeat tract may produce a suitable substrate (Fig 15). The incorporation of two unique restriction sites flanking the repeat region allows DNA replication products to be excised and quickly observed by gel electrophoresis, eliminating the need for sequencing to determine the composition of each product. A circular substrate would also allow loading of the T4 DNA polymerase accessory proteins so that the *in vitro* expansion assay could be even more *in vivo* like in terms of the replication apparatus. These suggestions could help give us more insight into the mechanism(s) of misalignment mutagenesis.

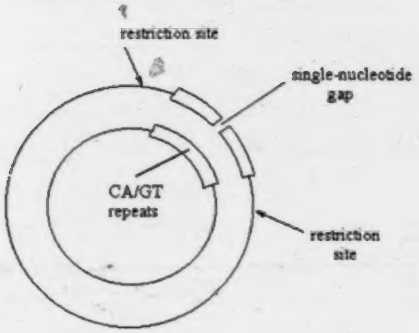


Figure 15. **The circular CA/GT repeat containing DNA substrate.** Future experiments involving the dinucleotide assay could make use of the substrate illustrated here. The DNA contains a CA/GT repeat sequence with a single-nucleotide gap within the repeat region. The CA/GT repeats are flanked on either side by restriction sites to allow for the removal of "expansion" products.

REFERENCES

Ashley, C. T., and S. T. Warren (1995) Trinucleotide repeat expansion and human disease. *Annu. Rev. Genetics* 29: 703-728.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and K. Struhl (1996) Current Protocols in Molecular Biology. John Wiley and Sons, Inc., Boston, MA.

Baker, R. P. and L. J. Reha-Krantz (1998) Identification of a transient excision intermediate at the crossroads between DNA polymerase extension and proofreading pathways. *Proc. Natl. Acad. Sci. U.S.A.* 95: 3507-3512.

Braithwaite, D. K., and J. Ito (1993) Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res.* 21: 787-802.

Brautigam, C. A., and T. A. Steitz (1998) Structural Principles for the Inhibition of the 3'→5' Exonuclease Activity of *Escherichia coli* DNA polymerase I by Phosphorothiaotes. *J. Mol. Biol.* 277: 363-367.

Clark, J. M. (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* 16: 9677-9686.

Clark, J. M., Joyce, C. M., and G. P. Beardsley (1987) Novel blunt-end addition reactions catalyzed by DNA polymerase I of Escherichia coli. *J. Mol. Biol.* 198: 123-127.

Dong, F., Weitzel, S. E., and P. H. von Hippel. (1996) A complex of T4 DNA replication helicase (gp41) and polymerase (gp43) can perform rapid and processive DNA strand-displacement synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 93: 14456-14461.

Drake, J. W., Allen, E. F., Forsberg, S. A., Preparata, R-M., and E. O. Greening (1969) Genetic control of mutation rates in bacteriophage T4. *Nature* 221: 1128-1132.

Duncan, J., Hamilton, L., and E. C. Friedberg (1976) Enzymatic degredation of uracil-containing DNA. *J. Virol.* 19: 338-345.

Eckstein, F. (1985) Nucleoside phosphorothiaotes. *Annu. Rev. Biochem.* 54: 367-402.

Frey, P. A. (1989) Chiral phosphorothioates: stereochemical analysis of enzymatic substitution at phosphorus. *Advan. Enzymol. Rel. Areas Mol. Biol.* 62: 119-201.

Frey, M. W., Nossal, N. G., Capson, T. L., and S. J. Benkovic (1993) Construction and characterization of a bacteriophage T4 DNA polymerase deficient in 3'→5' exonuclease activity. *Proc. Natl. Acad. Sci. U.S.A.* 90: 2579-2583.

Gura, T. (1997) New Kind of Cancer Mutation Found. *Science* 277: 1201-1202.

Hamada, H., Petrino, M. G., and T. Kakunga (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. U.S.A.* 79: 6465-6469.

Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D., and M. Perucho (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363: 558-561.

Joyce, C. M., and Steitz, T. A. (1995) Polymerase structures and function: variations on a theme. *J. Bacteriol.* 177: 6321-6329.

Kunkel, T. A. (1985a) The Mutational Specificity of DNA Polymerase-β during *in Vitro* DNA Synthesis. *J. Biol. Chem.* 260: 5787-5796.

Kunkel, T. A. (1985) The Mutational Specificity of DNA Polymerase-α and -γ during *in Vitro* DNA Synthesis. *J. Biol. Chem.* 260: 12866-12874.

Kunkel, T. A. (1990) Misalignment-mediated DNA synthesis errors. *Biochemistry* 29: 8003-8011.

Laken, S. J., Petersen, G. M., Gruber, S. B., Oddoux, C., Ostrer, H., Giardiello, F. M., Hamilton, S. R., Hampel, H., Markowitz, A., Klimstra, D., Jhanwar, S., Winawer, S., Offit, K., Luce, M. C., Kinzler, K. W., and B. Vogelstein (1997) Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nat. Genet.* 17: 79-83.

Maniatis, T., Fritsch, E. F., and J. Sambrook. (1982) Molecular Cloning A Laboratory Manual. Cold Spring Harbour Laboratory. New York.

Marquez, L. A. and L. J. Reha-Krantz (1996) Using 2-Aminopurine Fluorescence and Mutational Analysis to Demonstrate an Active Role of Bacteriophage T4 DNA Polymerase in Strand Separation Required for 3'→5' Exonuclease Activity. *J. Biol. Chem.* 271: 28903-28911.

Matsumoto, Y., and K. Kim (1995) Excisoin of deoxyribosephosphate residues by DNA polymerase beta during DNA repair. *Science* 269: 699-702.

Maxam, A. L., and W. Gilbert (1980) Sequencing End-Labeled DNA with Base-Specific Chemical Cleavages. *Methods in Enzymology* 65: 499-560.

Modrich, P. (1996) Mismatch repair in replication fidelity, genetic recombination and cancer biology. *Annu. Rev. Biochem.* 65: 101-133.

Mölling, K., Bolognesi, D. P., Bauer, H., Büsen, W., Plassmann, H. W., and P. Hausen (1971) Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nat. New. Biol.* 234: 240-243.

Muzyczka, N., Poland, R. L., and M. Bessman (1972) Studies on the biochemical basis of spontaneous mutation. I. A comparison of the deoxyribonucleic acid polymerases of mutator, antimutator, and wild type strains of bacteriophage T4. *J. Biol. Chem.* 247: 7116-7122.

Nossal, N. G. (1994) Chapter 5: The Bacteriophage T4 Replication Fork in The Molecular Biology of Bacteriophage T4. ASM Press: Washington D. C.

Peinado, M. A., Malkhosyan, S., Velazquez, A., and M. Perucho (1992) Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. *Proc. Natl. Acad. Sci. U.S.A.* 89: 10065-10069.

Petruska, J., Hartenstine, M. J., and M. F. Goodman (1998) Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats asscoiated with neurodegenerative disease. *J. Biol. Chem.* 273: 5204-5210.

Reha-Krantz, L. J. (1988) Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants. Relating structure to function. *J Mol. Biol.* 202: 711-724.

Reha-Krantz, L. J., and R. L. Nonay (1993) Genetic and biochemical studies of bacteriophage T4 DNA polymerase 3'→5' exonuclease activity. *J. Biol. Chem.* 268: 27100-27108.

Reha-Krantz, L. J., and R. L. Nonay (1994) Motif A of Bacteriophage T4 DNA Polymerase: Role in Primer Extension and DNA Replication Fidelity. *J. Biol. Chem.* 269: 5635-5643.

Reha-Krantz, L. J., Marquez, L. A., Elisseeva, E., Baker, R. P., Bloom, L. B., Dunford, H. B., and M. F. Goodman (1998) The proofreading pathway of bacteriophage T4 DNA polymerase. *J. Biol. Chem.* 273: 22969-22976.

Reha-Krantz, L. J., Nonay, R. L., and S. Stocki (1993) Bacteriophage T4 DNA Polymerase Mutations That Confer Sensitivity to the PP_i Analog Phosphonoacetic Acid. *J. Virol.* 67: 60-66.

Rubin, C. M., and C. W. Schmid (1980) Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucleic Acids Res.* 4: 4613-4619.

Schlötterer, C. and D. Tautz (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Res.* 20: 211-215.

Speyer, J. F., Karam, J. D., and A. B. Lenny (1966) On the role of DNA Polymerase in base selection. *Cold Spring Harbor Symp. Quant. Biol.* 31:693-697.

Spitzer, S., and Eckstein, F. (1988) Inhibition of deoxyribonucleases by phosphorothioate groups in oligodeoxynucleotides. *Nucleic Acids Res.* 16: 11691-11704.

Stocki, S. A., Nonay, R. L., and L. J. Reha-Krantz (1995) Dynamics of Bacteriophage T4 DNA Polymerase Function: Identification of Amino Acid Residues that Affect Switching between Polymerase and 3'→5' Exonuclease Activities. *J. Mol. Biol.* 254: 15-28.

Strand, M., Prolla, T. A., Liskay, R. M., and T. D. Petes (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature*. 365: 274-276.

Streisinger, G., Okada, Y., Emrich J., Newton J., Tsugita A., Terzaghi E., and M. Inouye (1966) Frameshift Mutations and the Genetic Code. *Cold Spring Harbour Symp. Quant. Biol.* 31: 77-84.

Tsurimoto, T., and Stillman, B. (1990) Functions of replication factor C and proliferating-cell nuclear antigen: Functional similarity of DNA polymerase accessory proteins from human cells and bacteriophage T4. *Proc. Natl. Acad. Sci. U.S.A.* 87: 1023-1027.

Wang, J., Yu, P., Lin, T. C., Konisberg, W. H., and T. A. Steitz (1996) Crystal structure of an NH_2-terminal fragment of T4 DNA polymerase and its complexes with single-stranded DNA and with divalent metal ions. *Biochemistry* 35: 8110-8119.

Whitehouse, H. L. K. (1963) A theory of crossing-over by means of hybrid deoxyribonucleic acid. *Nature* 199: 1034-1040.